基于正四棱锥结构的机器人声源定位系统研究 *

陈国良，徐　扬，黄晓琴

(武汉理工大学 机电工程学院，武汉 430070)

摘　要：声源定位成为机器人智能研究的重要方向。针对当前声源定位精度不理想、实时性不佳等问题，提出了一种正四棱锥麦克风阵列声源定位结构。采用时间延迟估计的声源定位方法，并提出时延值的快速搜索策略；推导了该结构的基于信号时延的时空映射关系，建立了声源目标位置的几何计算模型，并依据正四棱锥结构特点及冗余的时延值对值域划分，缩小求解范围，运用迭代算法得到声源的位置坐标，并通过双重筛选机制剔除错误的定位结果。实验结果证明了该结构及定位算法在提高系统定位精度和实时性能的有效性，能满足机器人应用中对声源定位的需求。

关键词：声源定位；四棱锥麦克风阵列；快速搜索策略；迭代算法；时延估计

中图分类号：TP391.42　　doi: 10.3969/j.issn.1001-3695.2017.11.0780

Sound source localization system for robot based on rectangular pyramid structure

Chen Guoliang, Xu Yang, Huang Xiaoqin

(*Institute of Electrical & Mechanical Engineering*, *Wuhan University of Technology*, *Wuhan* 430070, *China*)

Abstract: Sound source localization (SSL) has become an important field of intelligent robot research. This paper proposed a localization model based on the multi-microphone with the rectangular pyramid structure. Firstly, the system employed the time delay estimation (TDE) method based on the cross-power spectral phase algorithm and a fast search strategy of peak value to estimate the time delay differences. And then, the geometric equation based on the time-space mapping relationship is established to calculate the position of the sound source. For fast solving the equations, the multi-microphone array space is divided into several subspaces to narrow the solution range according to the characteristics of rectangular pyramid structure and the redundant time differences, and Newton iterative algorithm is introduced to solve the equations, which its solution with a large error is eliminated by the double screening mechanism. The experimental results show that the model and the algorithm are effective in improving the localization accuracy and real-time performance, which can meet the demand of sound source localization in robot application.

Key Words: sound source location; rectangular pyramid microphone array; fast search strategy; iterative algorithm; time delay estimation

0　引言

机器人作为当前科学技术领域的活跃因子，提高其智能化水平，实现与人类更加友好、便捷的交互是未来的发展趋势[1]。听觉是人体所具有的重要功能，也是人感知外界环境，并与外界环境进行交互的重要方式，再者听觉有其独特的优势，比如，听觉具有全向性，而且能够在黑暗的、有障碍物遮挡的环境下进行定位[2,3]。因此，基于听觉感知与定位的智能机器人技术得到了广泛的关注与快速的发展。

听觉定位技术在机器人领域具有广阔的应用前景[4]。一方面，从机器人服务于人的角度，该技术可使机器人定位说话人，并跟踪声源轨迹，提高人机交互的智能性和友好性；另一方面，从机器人识别环境的角度，机器人可以监听所在环境下的异常声音，并定位出声音源，找出声音事件发生的位置，便于机器人报警或者自行处理该异常事件。此外，还有护理机器人、智能雷达系统等重要行业领域的应用都表明了机器人听觉的积极作用，对人类的生产和生活有着较为深远的影响。

目前，常用的声源定位方法有三种：基于最大输出功率的可控波束形成方法[5,6]、高分辨率谱估计方法[7,8]和基于时延估计方法[9,10]。其中，基于时间延迟的定位方法，由于其计算量相对较小，易于实现而成为最普遍使用的方法[11]。文献[12]在机器人球形头部搭建仿双耳声源定位系统，为了降低基于信号相关的时延估计算法中，信号采样率对定位分辨率的影响，利用最大似然方法找出最大化互功率之和的声源轴向角，另外考虑

基金项目：国家自然科学基金资助项目（61373110 和 61672396）；武汉理工大学自主创新研究基金资助项目（175204004）

作者简介：陈国良（1972-），男，湖南娄底人，副教授，博士，主要研究方向为机器人技术（glchen@whut.edu.cn）；徐扬（1991-），男，河南驻马店人，硕士研究生，主要研究方向为机器人听觉技术；黄晓琴（1994-），女，湖北宜昌人，硕士研究生，主要研究方向为机器人听觉技术.

机器人头部带来的多径效应，降低外界干扰，利用基于 front-back 的多径补偿因子来修正时延估计。文献[13]设计了四麦克风的十字平面阵列，并提出一种基于时间差特征的空间栅格匹配算法，找到与待定位声源的时间差特征最匹配的栅格作为声源位置．该方法可以有效地避免几何定位方法的非线性方程组求解问题，复杂度较低．并且合理的麦克风阵列拓扑可以避免几何定位方法可能陷入局部最优点的问题，但阵列尺寸较大，限制了与机器人的结合。当前，优化麦克风阵列结构以提高与机器人的融合度；探索鲁棒的、定位精度高、对各种噪声广泛适用的抗噪声技术以适应人机交互中复杂的声学环境成为机器人听觉系统实用化过程中的研究热点。

本文基于声达时间差设计一种正四棱锥麦克风阵列定位系统。采用互功率谱相位算法获取时延，并提出峰值快速搜索策略；根据正四棱锥的结构特点和冗余的时延值对求解区域进行划分，缩小值域范围，最后采用牛顿迭代算法求得声源的最优化位置解，实验证明该模型和定位算法有效地提高了计算的速度和定位的精度。

1　正四棱锥声源定位原理

1.1　声源定位原理

同一声源传播到空间位置不同的两个麦克风m_1和m_2会产生时间差τ_{21}，τ_{21}由声源到达m_2的时间减去声源到达m_1的时间所得（同理，可以定义τ_{12}），这里设$\tau_{21}>0$，即声源接近于m_1，根据时间差可以得到声程差。由双曲线的性质可知，双曲线上任意一点到两个焦点距离差的绝对值为常数，即实轴长。据此可知，声源将在以这两个传声器的位置为焦点，声程差为实轴长的双曲线并且靠近m_1的分支线上。当有多个传声器的时候，就可以得到多个双曲线，则多个双曲线的交点即为声源的位置[14]。如图 1 所示，S为声源。

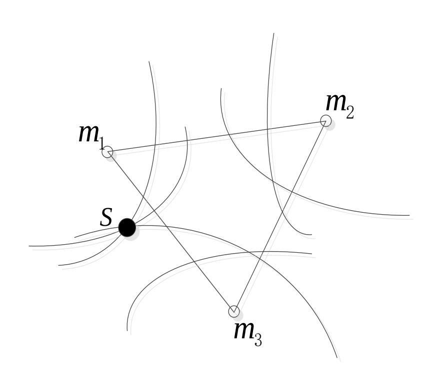


图 1　基于时延的双曲线声源定位原理图

1.2　时间延迟估计

设$x_1(t)$和$x_2(t)$分别为位置不同的两个麦克风接收到的同一声源的信号模型，则两信号模型必然存在相关性。互功率谱相位算法就是通过检测这两个信号的互相关函数的峰值点而得到声源到达两个麦克风的时延值[15,16]。其互相关函数可以表示为：

$$R_{x_1x_2}(\tau)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\varphi_{x_1x_2}(\omega)G_{x_1x_2}(\omega)e^{j\omega\tau}d\omega \tag{1}$$

$$\varphi_{x_1x_2}(\omega)=1/\left|G_{x_1x_2}(\omega)\right| \tag{2}$$

其中：$\varphi_{x_1x_2}$为加权函数；$G_{x_1x_2}(\omega)$为两麦克风获取信号$x_1(t)$和$x_2(t)$的互相关功率谱函数。$R_{x_1x_2}(\tau)$峰值所对应的横坐标即为时延值。

1.2.1 峰值的快速搜索策略

当信号的采样频率较高、数据量较大时，能缩小互相关函数的峰值搜索区间，快速准确地找出峰值点，也是提高算法性能和实时性的一种表现。

对于单个声源和两个麦克风而言，它们在同一平面上可存在两种几何关系。图 2 所示，一种是线性关系，即直线$s_2m_1m_2$和直线$s_3m_2m_1$，另一种是三角关系，即三角形$s_1m_1m_2$。

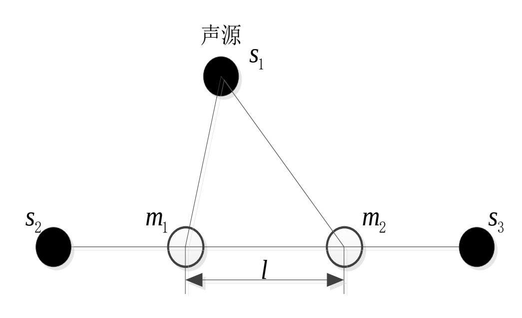


图 2　声源与麦克风的几何关系图

当两个麦克风的位置固定，距离间隔为l，并设声源到达麦克风m_1的距离与到达麦克风m_2的距离差值为d。由于三角形任意两边之差小于第三边，因此只有当声源与两麦克风共线，如声源位于图 2 中s_2或者s_3时，d取得最大值l和最小值$-l$。若声音的传播速度是v，采样频率为f，则易知峰值所在的采样点数区间为$[-l\times f/v, l\times f/v]$。如图 3 所示，虚线 1 和虚线 2 即为峰值搜索的边界。由此可以限定峰值的搜索范围，减少时延估计的耗时。

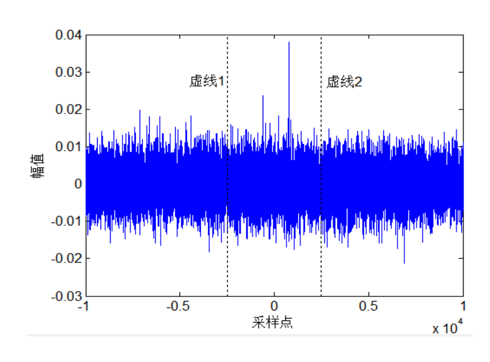


图 3　峰值搜索区间

1.2.2 时延值筛选

当所估计的时延值出现误差，且误差较大时，将导致最终定位的数值性错误，而且若在区域的分界线附近出现时延误差也极易产生定位的方向性错误。因此，在时延值快速搜索的基础上，更进一步，提出对时延值的筛选策略，其具体实现过程

如下：

a）理想状态下，在同一次时延估计中，麦克风i,j,k之间的时延值存在如下关系：

$$\tau_{ij}=\tau_{ik}+\tau_{kj} \tag{3}$$

例如$\tau_{12}=\tau_{13}+\tau_{32}$，但由于时延估计精度的问题，上式不可能处于完全相等的状态，它们之间有一个误差的存在。

b）针对 a）中所提到的误差设定一个合理的阈值ε，其中ε值可根据系统定位精度的需求来设定，并进行下面不等式的判断：

$$\left|\tau_{ij}-\left(\tau_{ik}+\tau_{kj}\right)\right|\le\varepsilon \tag{4}$$

c）当不满足该不等式时，说明时延估计的偏差较大，则摈弃该组时延值；当满足条件时，则该组时延通过筛选，进而进行定位计算。

1.3　基于时间延迟估计的正四棱锥声源定位

设计采用基于正四棱锥的三维麦克风阵列进行声源目标的定位。建立如图 4 所示的坐标系，m_0-m_4分别为五个麦克风且位于四棱锥的五个顶点之上，坐标分别为(x_i,y_i,z_i) $(i=0,1,2,3,4)$，待求声源点G的坐标为(x,y,z)，声音在空气中的传播速度为v_{sound}，声源点G到各个麦克风的距离分别为$s_i(i=0,1,2,3,4)$；则目标声源到达其他麦克风的距离与到达m_0的距离之差为

$$d_{i0}=s_i-s_0=v_{sound}\tau_{i0}\left(i=1,2,3,4\right) \tag{5}$$

其中τ_{i0}为声源到达m_i与到达m_0的时间差。

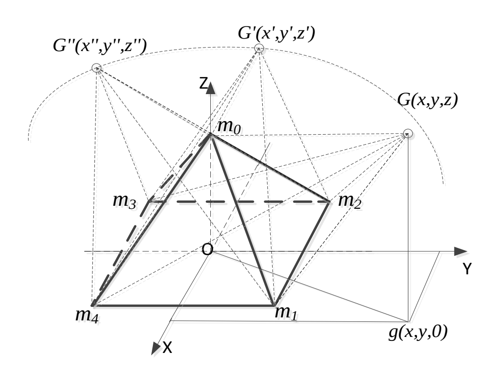


图 4　正四棱锥声源定位示意图

由图 4 的空间几何关系可建立声源定位数学模型如下：

$$\begin{cases} s_i=\sqrt{\left(x-x_i\right)^2+\left(y-y_i\right)^2+\left(z-z_i\right)^2} \\ s_0=\sqrt{\left(x-x_0\right)^2+\left(y-y_0\right)^2+\left(z-z_0\right)^2} \\ d_{i0}=s_i-s_0\left(i=1,2,3,4\right) \end{cases} \tag{6}$$

2　基于迭代优化的声源位置计算

方程组(6)中的任意三个方程联立，均可求解出声源的位置坐标。考虑到实际测量的误差，可以使其中三个方程分别联立，采用牛顿迭代优化算法求解出一组位置坐标，由组合原理易知可以得到四组解，再取四个解的平均值作为位置坐标的最终估计值。

2.1　分区迭代过程

初始位置坐标的选取对牛顿迭代算法的收敛速度以及最终的收敛结果产生较大的影响。为了初步确定声源位置坐标范围从而减小迭代次数，也使每个取值范围内具有唯一的优化点，由正四棱锥的结构特点可知麦克风阵列在定位时会产生冗余的时间差，例如$\tau_{i1}(i=2,3,4)$，当$\tau_{21}\ge0$、$\tau_{31}\ge0$、$\tau_{41}\ge0$且$\tau_{21}\ge\tau_{41}$时，可知声源的位置坐标满足$\{(x,y,z)|x\ge0,y\ge0\&x\ge y\}$，方位角也必然在$[0^\circ\ 45^\circ]$的范围内。因此，可以充分利用这些冗余时间差进行求解区间的划分，照此规则，如图 5 所示，按45°为一个区域，将定位空间等划分为八个区域。当得知麦克风阵列测得的时间差以及时间差之间的相对大小关系时，即可初步将声源的位置坐标锁定在其中一个区域内，由此可以缩小求解的范围区间，加快计算速度。

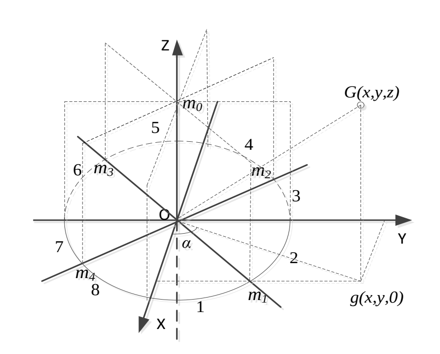


图 5　定位区间划分示意图

由牛顿迭代算法原理，并结合阵列定位数学模型，将(6)式转换为四组标准的迭代表达式，其中一组为

$$\begin{aligned} f_i\left(x,y,z\right)=&\sqrt{\left(x-x_i\right)^2+\left(y-y_i\right)^2+\left(z-z_i\right)^2} \\ &-\sqrt{\left(x-x_0\right)^2+\left(y-y_0\right)^2+\left(z-z_0\right)^2}-d_{i0}\left(i=1,2,3\right) \end{aligned} \tag{7}$$

其 Jacobi 矩阵为

$$J=\begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} \\ \frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} & \frac{\partial f_3}{\partial z} \end{bmatrix} \tag{8}$$

则声源坐标位置可以表示为

$$\begin{bmatrix} x^{(k+1)} \\ y^{(k+1)} \\ z^{(k+1)} \end{bmatrix}=\begin{bmatrix} x^{(k)} \\ y^{(k)} \\ z^{(k)} \end{bmatrix}-J^{-1}\begin{bmatrix} f_1\left(x^{(k)},y^{(k)},z^{(k)}\right) \\ f_2\left(x^{(k)},y^{(k)},z^{(k)}\right) \\ f_3\left(x^{(k)},y^{(k)},z^{(k)}\right) \end{bmatrix} \tag{9}$$

2.2　定位结果筛选

通过上节对时延值的筛选，为迭代运算进行准确的区间划分和求解提供了一定的保障。当迭代运算在正确的区间运行时，四组定位结果较为准确，且彼此之间的误差较小；然而当声源极其接近分界线时，很小的时延误差也会造成求解区域的判断

错误，在错误的区间进行迭代运算，将得到错误的定位结果，且定位结果之间误差较大，甚至无法得出结果。因此，可以利用这一差异，采用与时延值筛选相似的方法对误差设定阈值$\delta_k(k=1,2,3)$，并设四组定位结果分别为(x_i,y_i,z_i) $(i=1,2,3,4)$，则：

$$\begin{cases}|x_i-x_j|\le\delta_1\\|y_i-y_j|\le\delta_2\\|z_i-z_j|\le\delta_3\end{cases}\quad(10)$$

其中：$i,j=1,2,3,4$且$i\ne j$，δ_k值根据系统定位精度需求而定。

当定位结果满足上面三个不等式时，则通过定位筛选，认为定位结果是正确的，否则予以剔除。

将通过筛选机制得到的四组解取平均值作为最终的定位结果(x,y,z)，并由此可以求解出声源目标到原点距离R、方位角α、仰角β。

$$R=\sqrt{x^2+y^2+z^2}\quad(11)$$

$$\alpha=\arctan\frac{y}{x}\quad(12)$$

$$\beta=\arctan\frac{\sqrt{x^2+y^2}}{z}\quad(13)$$

3 实验结果与分析

3.1 实验样机

图 6 为声源定位的测试样机。样机在硬件上主要包括正四棱锥阵列、麦克风模块、数据采集卡以及计算机等。正四棱锥阵列设计后由 3D 打印完成，棱长为 25cm，在四棱锥的五个顶点处设计有麦克风安装孔，可实现麦克风与阵列模型较为稳固的机械连接，麦克风采用 DGO-6050CD-P，为全向性驻极体麦克风，频率范围为 20Hz-16KHz。麦克风的具体参数如表 1 所示。样机实验所用的移动机器人平台为“旅行家二号”轮式移动机器人。

表 1 麦克风技术参数

参数	数值
尺寸	9.7X6.7mm
灵敏度	-48-66dB
方向特征	全指向
工作电压范围	1.5-10V
阻抗	低阻抗
电流消耗	最大500μA

选用 USB_HRF4626 数据采集卡完成对声音信号的采集与传输，其一端通过 USB 与计算机连接，另一端通过 5 根数据线与麦克风阵列相连，如图 7 所示。该款数据采集卡是一种基于 USB 总线的高速高精度同步数据采集卡，可以很方便地与外部设备连接；其具有 8 路 16 位高精度同步 AD 采集通道，本实验使用其中的 5 路通道即可。单通道的采样频率最大值达到 50KHz。其具体参数如表 2 所示。



图 6 麦克风阵列实物图

表 2 数据采集卡技术参数

参数	数值
AD类型	双极性
精度	16bit
输入电压量程	-10-10V
电压通道数	8
最大耐压值	±16.5V
最高采样频率	50KHz



图 7 数据采集卡的连接图

各麦克风获取声源信号并经过预处理后，采用互功率谱相位算法获取两两麦克风的时延值。例如对于麦克风m_i和m_j，将两麦克风接收到的声音信号通过傅里叶变换从时域变换到频域，在频域内通过互功率谱加权函数突出声源信号抑制噪声影响，再将信号变换到时间域，最后通过检测这两个信号的互相关函数的峰值点而得到时延值τ_{ij}，如图 8 所示。

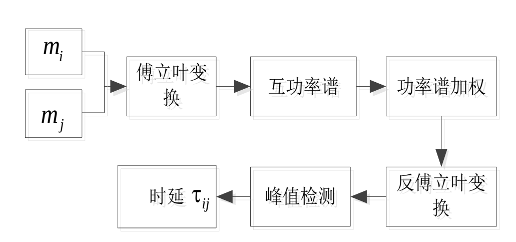


图 8 时延获取的信号处理过程图

编写上位机程序求取麦克风的各时延值，然后采用第二节的迭代优化定位算法即可得出声源的位置坐标，其上位机界面

如图 9 所示。

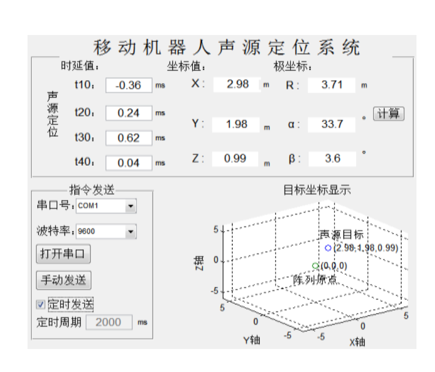


图 9　上位机界面



图 10　实验环境

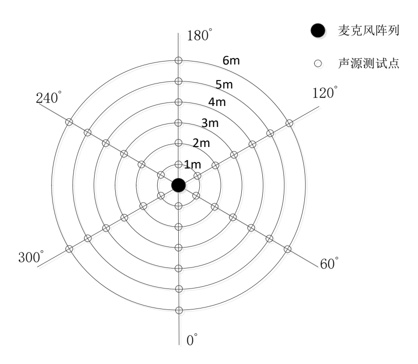


图 11　测量点分布示意图

3.2　实验分析

通过具体实验来验证所建麦克风阵列模型以及定位算法性能的优劣。该实验被安排在一间较为宽敞的实验室内，可保证在室内任何一个角落都可以接收到声音信号。实验环境相对安静，但也存在一定噪声的干扰，噪声的来源主要是室内照明设备的声响以及电脑的散热，实验环境如图 10 所示。定位阵列模型被安装在机器人平台上并建立起通信连接，声源选用已录制好的一段拍手的声音信号，因为拍手声的信号能量集中，且无声段与有声段区分明显，易于实验的进行，声音信号大小为 45 dB。样机采样频率设置为 50 kHz。

实验总共进行了36个位置的测量，声源设备到阵列中心点的距离R分别为 1m、2m、3m、4m、5m、6m，每个距离点上方位角α分别为0°、60°、120°、180°、240°、300°，仰角β均为30°，如图 11 所示。

实验采取的数据处理流程为：信号预处理阶段，信号分帧帧长为 1024 个采样点，帧移为 512 个采样点；时延估计阶段，采用互功率谱相位算法进行时延估计，并通过筛选机制剔除误差较大的时延估计值；定位阶段，采用所提出的分区迭代方法求取目标定位结果，并对定位结果进行筛选。实验在每个测量点重复做 5 次测量，然后取平均值作为该测量点的最终定位结果。得到的实验结果如表 3 所示。

表 3　定位实验结果统计表　　角度单位：（度）　距离单位：（m）

真实值（α，β，R）	实验值（α，β，R）	真实值（α，β，R）	实验值（α，β，R）	真实值（α，β，R）	实验值（α，β，R）
(0,30,1)	(0.9,30.8,1.05)	(60,30,1)	(59.1,29,7,1.06)	(120,30,1)	(120.7,30.6,0.91)
(0,30,2)	(1.1,30.3,2.08)	(60,30,2)	(60.6,30.9,2.10)	(120,30,2)	(118.9,28.9,2.08)
(0,30,3)	(1.0,29.8,2.88)	(60,30,3)	(60.8,30.6,3.13)	(120,30,3)	(121.1,30.4,3.12)
(0,30,4)	(0.8,30.5,3.85)	(60,30,4)	(60.9,29.3,3.88)	(120,30,4)	(120.7,30.4,4.13)
(0,30,5)	(0.9,30.5,5.18)	(60,30,5)	(60.9,29.1,5.19)	(120,30,5)	(119.0,29.7,4.86)
(0,30,6)	(1.0,30.4,6.22)	(60,30,6)	(59.2,30.0,6.19)	(120,30,6)	(121.0,29.3,6.21)
真实值（α，β，R）	实验值（α，β，R）	真实值（α，β，R）	实验值（α，β，R）	真实值（α，β，R）	实验值（α，β，R）
(180,30,1)	(181.1,30.6,1.07)	(240,30,1)	(239.0,30.2,1.04)	(300,30,1)	(301.3,29.5,1.07)
(180,30,2)	(181.2,30.8,1.88)	(240,30,2)	(240.9,30.2,2.11)	(300,30,2)	(300.8,30.6,1.90)
(180,30,3)	(180.9,30.9,3.09)	(240,30,3)	(241.4,29.3,2.85)	(300,30,3)	(301.3,29.3,3.12)
(180,30,4)	(181.3,29.4,4.16)	(240,30,4)	(241.2,30.9,3.87)	(300,30,4)	(298.8,29.6,4.14)
(180,30,5)	(180.9,29.6,5.18)	(240,30,5)	(241.0,30.6,5.17)	(300,30,5)	(301.4,29.3,5.16)
(180,30,6)	(181.3,30.3,6.16)	(240,30,6)	(241.4,30.9,5.82)	(300,30,6)	(301.2,30.7,6.24)

由表 3 的定位结果，采用均方根误差(RMSE)作为实验定位精度的判断标准。由于本实验着重实现机器人对室内声源目标的定位，仰角的影响因素相对较小。因此，本实验将重点分析定位距离与定位方位角的精度问题，暂不考虑仰角信息（由表

3 可知仰角误差在1.2°以内），由此可得到定位误差的曲线图，如图 12 到图 15 所示。

由图 12、13 可知，在同一方位角度下，随着声源与定位阵列距离的增大，距离定位误差与角度定位误差呈现出不同的关系曲线。即定位距离误差随着距离的增大而增大，在6m的范围内，最大距离定位误差保持在0.25 m以内，最小误差只有0.05m左右；而方位角定位误差并没有随着声源距离的变化而出现显著的变动，误差在1.5°以内，保持了较好的稳定性，定位精度相对较高。从距离误差与角度误差的相对大小来看，角度定位精度要优于距离定位精度。初步分析，造成这种定位精度差异的重要原因是距离的计算式（11）是由乘方与加法组成，这也会造成误差的累积；而角度值的求解公式恰恰相反，是由除式组成的，这对误差的消除起到了一定的作用。

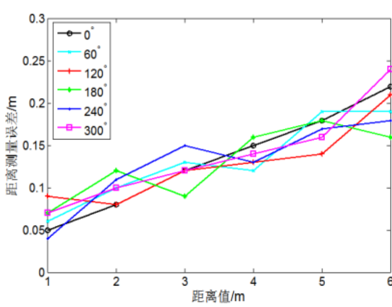


图 12　定位距离误差与距离的关系曲线

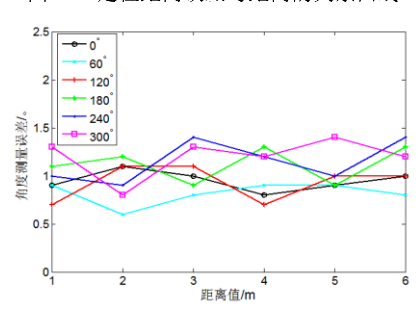


图 13　定位角度误差与距离的关系曲线

由图 14 和图 15 可知，在同一距离的情况下，随着方位角的变化，距离定位误差与方位角定位误差均呈现出较平稳的关系曲线。距离定位误差的波动范围保持在 0.08m 以内，方位角定位误差的波动范围维持在0.6°以内。从阵列结构上分析可知，由于正四棱锥的对称性，使得从各个角度接收到声源信号较为均衡。因此距离定位误差和角度定位误差并没有随着方位角的不同而出现显著的波动。

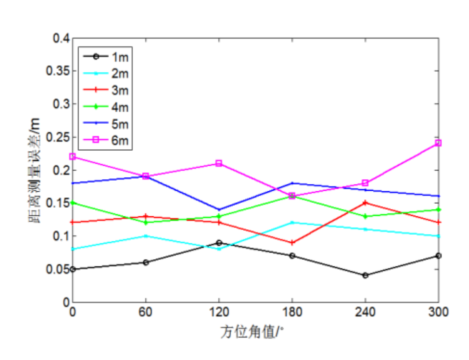


图 14　定位距离误差与方位角的关系曲线

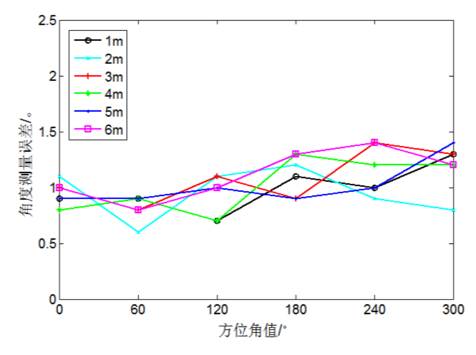


图 15　定位角度误差与方位角的关系曲线

3.3　相对误差分析

图 16 和图 17 分别表示定位距离相对误差与定位方位角相对误差随距离与角度的变化曲线。从图中可以看出，两者的相对误差均随着两者的增大而出现减小的趋势；在设定的实验距离范围内，定位距离的相对误差最大值为，最小值仅为 2.67%，方位角的相对误差则在　以内，能较好地满足室内目标定位需求。

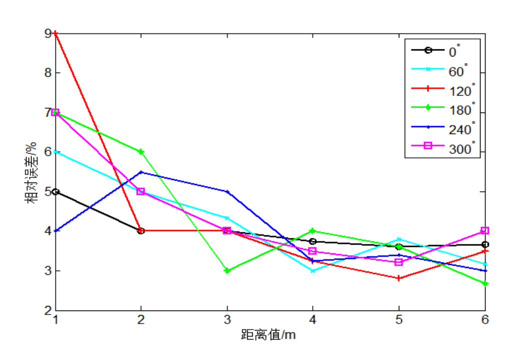


图 16　定位距离的相对误差

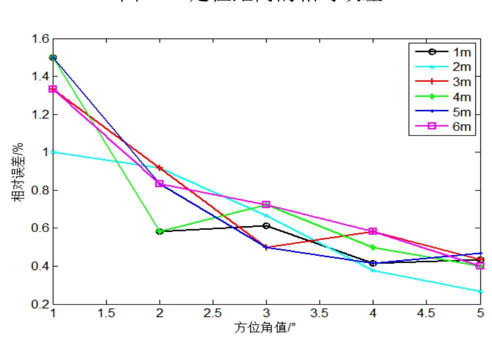


图 17　定位角度的相对误差

3.4　性能对比验证

科大讯飞作为智能语音技术的先行者，在语音识别、声源定位等多项技术上拥有领先的成果。旗下发布的新一代平面六麦克风环形阵列 XFR801，集成声源定位、语音增强、语音唤醒、语音打断等多种功能，广泛应用于机器人、智能家居、智能硬件等多种场景，其算法复杂度较高。而声源定位作为本文的研究重点和研究目标，因此重点从声源定位的角度将本阵列与 XFR801 进行比较，其结果如表 4 所示。有关 XFR801 的其他功能与技术是本文今后需要学习和研究的方向。

表 4　性能对比结果

阵列	定位角度范围	水平角度定位误差	是否具有仰角定位	是否具有距离定位	是否具有语音增强	是否具有语音唤醒
XFR801	360°	≤10°	否	否	是	是
测试样机	360°	≤1.5°	是	是	否	否

4　结束语

设计实现了基于正四棱锥的空间声源定位系统。该系统采用互功率谱相位算法获取时延，并制定了时延值快速搜索策略；建立起合适的声源定位数学模型，并对求解区域进行划分，初步确定定位范围，采用牛顿迭代算法求取声源的位置坐标，并利用双重筛选机制剔除错误的定位结果。从实验结果与分析可以看出，该模型和算法达到了较好的定位效果，在一定程度上弥补了先前声源定位模型定位精度不理想的问题，可以满足实际的应用需求。进一步，将该装置与移动机器人结合，可实现对声源目标的自主定位与跟踪。

参考文献：

[1] 任福继，孙晓．智能机器人的现状及发展 [J]. 科技导报, 2015, 33 (21): 32-38.

[2] Marinescu R S, Buzo A, Cucu H, et al. Fast accurate time delay estimation based on enhanced accumulated Cross-power Spectrum Phase [C]// Proc of Signal Processing Conference. 2014: 1-5.

[3] Mayyas K, Momani F. An LMS adaptive algorithm with a new step-size control equation [J]. Journal of the Franklin Institute, 2011, 348 (4): 589-605.

[4] Valin J M, Michaud F, Rouat J, et al. Robust sound source localization using a microphone array on a mobile robot [C]// Proc of IEEE//RSJ International Conference on Intelligent Robots and Systems. 2016: 1228-1233.

[5] Brutti A. Tracking of multidimensional TDOA for multiple sound sources with distributed microphone pairs [J]. Computer Speech & Language, 2013, 27 (3): 660—682.

[6] Baumgartner R, Majdak P. Assessment of Sagittal-plane Sound Location Performance in Spatial-audio [M]. Berlin: Springer, 2013: 93-119.

[7] Joshua M Sachar, Harvey F Silverman. Microphone Position and Gain Calibration for a Large-Aperture Microphone Array [J]. IEEE Transactions on Speech and Audio Processing, 2005, 13 (1): 42-52.

[8] Salvati D, Canazza S. Adaptive time delay estimation using filter length constraints for source localization in reverberant acoustic environments [J]. IEEE Signal Processing Letters, 2013, 20 (5): 507-510.

[9] 刘慧珠．基于麦克风阵列的声源方位估计算法研究 [D]. 成都：西南交通大学, 2015.

[10] Pourmohammad A, Ahadi S M. The real time high accuracy 3-d phat-based sound source localization using a simple 4-microphone arrangement [J]. IEEE Systems Journal, 2012, 6 (3): 455-468.

[11] Tsiami A, Katsamanis A, Maragos P, et al. Experiments in acoustic source localization using sparse arrays in adverse indoors environments [C]// Proc of Signal Processing Conference. 2014: 2390-2394.

[12] Kim U H, Mizumoto T, Ogata T, et al. Improvement of speaker localization by considering multipath interference of sound wave for binaural robot audition [C]// Proc of IEEE//RSJ International Conference on Intelligent Robots and Systems. 2011: 2910-2915.

[13] Li X, Liu H, Yang X, et al. Sound source localization for mobile robot based on time difference feature and space grid matching [C]// Proc of IEEE//RSJ International Conference on Intelligent Robots and Systems. 2011: 2879-2886.

[14] Nesta F, Omologo M. Generalized state coherence transform for multidimensional TDOA estimation of multiple sources [J]. IEEE Trans on Speech Processing, 2012, 20 (1): 246-260.

[15] 张雷岳，张兴敢，刘超．麦克风阵列声源定位中时延估计的改进 [J]. 南京大学学报：自然科学版, 2015, 51 (1): 25-30.

[16] 孙昊，仲维灿，刘贺洋．正四面体麦克风阵列声源定位模型研究 [J]. 计算机真, 2015, 32 (2): 378-382.